

University of Wollongong  
**Research Online**

---

Faculty of Informatics - Papers (Archive)

Faculty of Engineering and Information  
Sciences

---

2009

## Ranking method for optimizing precision/recall of content-based image retrieval

Jun Zhang

*University of Wollongong*, [jz484@uow.edu.au](mailto:jz484@uow.edu.au)

Lei Ye

*University of Wollongong*, [lei@uow.edu.au](mailto:lei@uow.edu.au)

Follow this and additional works at: <https://ro.uow.edu.au/infopapers>



Part of the [Physical Sciences and Mathematics Commons](#)

---

### Recommended Citation

Zhang, Jun and Ye, Lei: Ranking method for optimizing precision/recall of content-based image retrieval 2009.

<https://ro.uow.edu.au/infopapers/3361>

Research Online is the open access institutional repository for the University of Wollongong. For further information contact the UOW Library: [research-pubs@uow.edu.au](mailto:research-pubs@uow.edu.au)

---

## Ranking method for optimizing precision/recall of content-based image retrieval

### Abstract

The ranking method is a key element of content-based image retrieval (CBIR) system, which can affect the final retrieval performance. In the literature, previous ranking methods based on either distance or probability do not explicitly relate to precision and recall, which are normally used to evaluate the performance of CBIR systems. In this paper, a novel ranking method based on relative density is proposed to improve the probability based approach by ranking images in the class. The proposed method can achieve optimal precision and recall. The experiments conducted on a large photographic collection show significant improvements of retrieval performance.

### Disciplines

Physical Sciences and Mathematics

### Publication Details

Zhang, J. & Ye, L. (2009). Ranking method for optimizing precision/recall of content-based image retrieval. UIC-ATC 2009 - Symposia and Workshops on Ubiquitous, Autonomic and Trusted Computing in Conjunction with the UIC'09 and ATC'09 Conferences (pp. 356-361). California, USA: The Institute of Electrical and Electronics Engineers, Inc..

# Ranking Method for Optimizing Precision/recall of Content-based Image Retrieval

Jun Zhang and Lei Ye

*School of Computer Science and Software Engineering,  
University of Wollongong,  
Wollongong NSW 2522 Australia  
Email: {jz484, lei}@uow.edu.au*

## Abstract

*The ranking method is a key element of Content-based Image Retrieval (CBIR) system, which can affect the final retrieval performance. In the literature, previous ranking methods based on either distance or probability do not explicitly relate to precision and recall, which are normally used to evaluate the performance of CBIR systems. In this paper, a novel ranking method based on relative density is proposed to improve the probability based approach by ranking images in the class. The proposed method can achieve optimal precision and recall. The experiments conducted on a large photographic collection show significant improvements of retrieval performance.*

## 1. Introduction

Content-based image retrieval (CBIR) is an active topic in information technique research area [1], [2] due to its wide applications, e.g., digital museum management, image search on web and private album organization. Comparison with traditional text-based retrieval paradigm, one advantage of CBIR is its ability to search by example, e.g., normally users are allowed to express his query with some example images. In CBIR systems, the content of images are characterized using their visual features, such as MPEG-7 visual descriptors [3]. Images are retrieved from the collection in accordance of their similarities to the content of user-provided query images.

As the retrieval result of a CBIR system is an ranked image sequence, the ranking method is an key element for the final retrieval performance. In previous work, the common ranking methods can be categorized according to the similarity measures as follows.

Distance based ranking methods: to minimize the average distances between retrieved images and the query image [4]–[6]. The image similarity is characterized by feature distances between the query image and images in the collection. Then all images in collection will be ranked according to their similarities and top ranked  $k$  images in the image sequence will be returned and displayed. The advantage of the feature distance based approach is the reduced computation complexity.

Probability based ranking methods: to minimize the average error probability that the retrieved images are relevant to the query image [7], [8]. In which the image similarity is represented using the posterior probability based on probabilistic models. And top  $k$  images with maximal posterior probability will be returned and displayed. The advantage of this approach is its significant gain in retrieval accuracy.

In [7], N. Vasconcelos treats image retrieval as a probabilistic classification problem and solved it using the approach of maximum a posteriori probability (MAP). In particular, the image classes are ranked in accordance of the posterior probability that the query image belongs to the image classes. However, the ranking of images in the class is not addressed in the classification approach.

As we know, the evaluation of the image retrieval performance is still an open problem [9]. Precision and recall, as the standard performance criteria in the information retrieval, are still the most popular performance measure in CBIR, so we will use them to evaluate the retrieval performance in this paper. From the view of performance evaluation, existing image retrieval approaches with different ranking methods may achieve their own optimizing objectives, but generally not the precision/recall.

These observations motivate the work presented in this paper. We argue that the order of images determines the retrieval performance in terms of precision and recall and the probability based ranking method presented in [7] should be improved by considering the ranking images in the class. In this paper, the image retrieval is treated as a classification problem. Precision/recall graph will be used as performance criteria, which is different to the probability of error for the conventional classification problem [10]. We use the same assumption in [7] that the image classes have been set up before doing image retrieval. In our CBIR system, the image retrieval has two steps, First, the image classes are ranked in accordance of the posterior probability that the query image is belonging to. This is achieved in [7]. Second, the images in classes are ranked for optimizing the precision and recall. This will be addressed in this paper. To achieve the goal, we start from addressing the problem that how to rank the next image for optimizing the precision/recall and show the new ranking method matching the MAP criteria. Then, we will explore that the proposed ranking method is able to optimize

the precision/recall in class.

The rest of paper is organized as follows. Section 2 gives the problem formulation and the new ranking method. How to achieve the optimal precision and recall using the proposed ranking method will be explored in Section 3. Section 4 reports the result of experiments for evaluating the retrieval performance of the new ranking method. Section 5 concludes the paper.

## 2. Improved ranking method

In this section, image retrieval will be treated as a classification problem and an improved probability based ranking method will be proposed by considering ranking images in class.

### 2.1. Problem formulation

The content of an image is described using visual features. We denote the high dimension feature space by  $\mathbf{X}$ , and  $\mathbf{x}$  and  $\mathbf{q}$  are the feature vectors of an image in collection and the feature vector of a query image, respectively. As in [7], we assume that the image classes in collection are existing, i.e.,  $\Omega = \{\omega_1, \dots, \omega_M\}$ . Then the image retrieval problem consists of two sub-problems.

The first one is how to rank image classes, which has been addressed as a conventional classification problem [7], i.e., how to categorize the query image  $\mathbf{q}$  into  $M$  image classes. Considering in image retrieval application user always just check a few retrieval result, e.g., 200 images, one strategy is to choose only one class and display images in the class. In this case, under MAP criteria the image class with highest posterior probability  $P(\omega_i|\mathbf{q})$  should be chosen. The other strategy is to rank all image classes and display the images in classes one by one. In this case, the posterior probability will be used as the similarity between the query image and the image classes.

The other one is how to rank images in class. Since in collection one image  $\mathbf{x}$  is categorized to  $\omega_i$  with the probability of error  $1 - P(\omega_i|\mathbf{x})$ , the order of images in the class will affect the retrieval performance in terms of precision and recall. While the problem is not considered in [7], we will address it in next subsection.

### 2.2. Ranking images in class

In this subsection, we address the problem that how to rank images in the class for optimizing the precision and recall. Precision  $P_r$  and recall  $R_e$  are normally used for evaluate the retrieval performance in CBIR, which are defined as

$$\begin{aligned} P_r &= \frac{\#(\text{retrieved relevant images})}{\#(\text{retrieved images})}; \\ R_e &= \frac{\#(\text{retrieved relevant images})}{\#(\text{relevant images in collection})}, \end{aligned} \quad (1)$$

where  $\#$  means the number of components. Given the number of retrieved images  $k$  and the number of relevant images in collection  $N_r$ ,  $P_r$  and  $R_e$  can be expressed as

$$P_r(R_e) = \frac{N_r}{k} R_e. \quad (2)$$

This is not a linear function, in which  $N_r$  can be considered as a constant, but  $R_e$  increasing will lead to  $k$  increasing. We will leave the hidden relationship between  $R_e$  and  $k$  and just use  $P_r(R_e)$  to analyze the retrieval performance qualitatively.

We start to consider the problem that how to rank next one image for maximizing the precision and recall, under the assumptions as following

- Setting up the image classes and ranking image classes are independent.
- If the query image and the collection image belong to the same image class, they are considered to be relevant.

So the query image  $\mathbf{q}$  and the image  $\mathbf{x}$  in class  $\omega_i$  are relevant with the probability  $P(\omega_i|\mathbf{x}) \cdot P(\omega_i|\mathbf{q})$ , where  $P(\omega_i|\mathbf{x})$  is the probability that  $\mathbf{x}$  belongs to the class  $\omega_i$  and  $P(\omega_i|\mathbf{q})$  is the probability that  $\mathbf{q}$  belongs to the class  $\omega_i$ . In the class  $\omega_i$ , let the precision and recall are  $P_{r,j}$  and  $R_{e,j}$ , respectively, before ranking the  $(j+1)$ th image. After ranking the  $(j+1)$ th image  $\mathbf{x}$  the precision and recall are

$$\begin{aligned} P_{r(j+1)} &= \frac{j \cdot P_{r,j} + 1 \cdot P(\omega_i|\mathbf{x}) \cdot P(\omega_i|\mathbf{q})}{j+1}; \\ R_{e(j+1)} &= \frac{N_r \cdot R_{e,j} + 1 \cdot P(\omega_i|\mathbf{x}) \cdot P(\omega_i|\mathbf{q})}{N_r}. \end{aligned} \quad (3)$$

To maximize the precision  $P_{r(j+1)}$  and recall  $R_{e(j+1)}$  simultaneously, the  $(j+1)$ th image should be

$$\mathbf{x} = \arg \max_{\mathbf{x}} P(\omega_i|\mathbf{x}). \quad (4)$$

Obviously, the above function is MAP criteria. So in this case, optimal precision and recall match MAP criteria.

### 2.3. A novel ranking method

Considering in practical applications, the high dimension of feature vector is a problem and distance metric is a popular solution. We try to rank distances instead of feature vectors, which can be looked as the extension of the approaches presented in [11], [12]. Given a query image  $\mathbf{q}$ , all images are mapped into distances by a transformation  $D: \mathbf{X} \rightarrow \mathbf{T}$ , and  $t = D(\mathbf{x}, \mathbf{q})$  is a distance in  $\mathbf{T}$ . As following, we use a distance to refer to an image in the collection. With the distance  $t$  between  $\mathbf{x}$  and  $\mathbf{q}$ , (4) can be rewrite as

$$t = \arg \max_t P(\omega_i|t). \quad (5)$$

Using Bayes formula

$$\begin{aligned} t &= \arg \max_t \frac{p(t|\omega_i) P(\omega_i)}{p(t)} \\ &= \arg \max_t \frac{p(t|\omega_i)}{p(t)}. \end{aligned} \quad (6)$$

We introduce the *relative density* of the distance of relevant images to that of images in the collection, that is,

$$f_r(t) = \frac{p(t|\omega_i)}{p(t)}. \quad (7)$$

We argue that ranking images in class in accordance of their relative density based on feature distance can achieve optimal precision and recall, which will be carefully checked in Section 3.

## 2.4. Discussion

In this Section, we explore the relationship between the proposed ranking method and conventional ranking method based on feature distance. We assume that the distribution of collection images in the distance space is uniform. Such that ranking images in accordance of their  $f_r(t)$  is equivalent to ranking images in accordance of their  $p(t|\omega_i)$ .

First we consider Gaussian model for  $p(t|\omega_i)$ , that is,

$$\begin{aligned} p(t|\omega_r) &= G(t, \mu, \sigma) \\ &= \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(t-\mu)^2}{2\sigma^2}}, \quad t \geq 0 \end{aligned} \quad (8)$$

where  $\mu$  is mean and  $\sigma$  is variance. Let the centroid of ground truth is  $\mathbf{x}_c$ , then we have

$$\mu = D(\mathbf{x}_c, \mathbf{q}). \quad (9)$$

If  $\mu = 0$ , i.e., the query image is the centroid of the class, then ranking images according to  $p(t|\omega_r)$  is equivalent to ranking images according to  $t$  which is the conventional ranking method based on feature distances. In other words, the conventional ranking method can achieve optimal precision and recall when the query image is the centroid of the class. If  $\mu > 0$ , i.e., the query image is not the centroid of the class, then ranking images according to  $p(t|\omega_r)$  is equivalent to ranking images according to  $|t - \mu|$ . In this case conventional ranking method based on feature distances can not achieve the optimal precision and recall though it has the least average feature distance. So if we assume the distribution of the class is Gaussian, we should try to estimate the centroid of the class and rank images from the centroid.

In practical applications, if more empirical distribution model is considered for  $p(t|\omega_i)$ , which can be trained using the existing image classes and the query image. Another possible way is to estimate the distribution parameters using multiple images, e.g., in relevance feedback.

## 3. Optimal precision/recall

In this section, we explain why the proposed ranking method can achieve optimal precision and recall from the global view. To compare two precision/recall graphs, we need to define a method of comparison.

**Definition 1.** *The retrieval performance of one ranking method  $Pr_a(Re)$  is strongly better than that of the other ranking method  $Pr_b(Re)$ , if and only if*

$$Pr_a(Re) \geq Pr_b(Re); \quad \forall Re. \quad (10)$$

In accordance of Definition 1, we have that *best optimal ranking* of the retrieved images is the ranking that is strongly better than any rankings.

The property of strongly better rankings is transitive. From the definition, we have

$$\begin{aligned} Pr_a(Re) &\geq Pr_b(Re); \quad \forall Re, \\ Pr_b(Re) &\geq Pr_c(Re); \quad \forall Re. \end{aligned} \quad (11)$$

So

$$Pr_a(Re) \geq Pr_c(Re); \quad \forall Re. \quad (12)$$

That is,  $Pr_a(Re)$  is strongly better than  $Pr_c(Re)$ .

Image retrieval produces an ranked image sequence. For convenience, we refer the performance of a retrieval method or a ranking method as the performance of the ranked image sequence. And the retrieval performance is evaluated using the graph  $Pr(Re)$ .

For two image sequences

$$\begin{aligned} \chi_a &:= (\mathbf{x}_1, \dots, \mathbf{x}_{j-1}, \mathbf{x}_j, \mathbf{x}_{j+1}, \mathbf{x}_{i+2}, \dots), \\ \chi_b &:= (\mathbf{x}_1, \dots, \mathbf{x}_{j-1}, \mathbf{x}_{j+1}, \mathbf{x}_j, \mathbf{x}_{i+2}, \dots), \end{aligned}$$

$f_{r,j}$  is the relative density of the distance between  $\mathbf{x}_j$  and  $\mathbf{q}$ ,  $\forall \mathbf{x}_j \in \omega_i$  and  $f_{r,j} \geq f_{r,j+1}$ . We analyze the retrieval performance of these two sequences as following.

It's clear that the difference between  $\chi_a$  and  $\chi_b$  is the order of  $\mathbf{x}_j$  and  $\mathbf{x}_{j+1}$ . Let  $Pr = P_0$ ,  $Re = R_0$ , after ranking  $\mathbf{x}_{j-1}$ , and  $Pr = P_3$ ,  $Re = R_3$ , before ranking  $\mathbf{x}_{j+2}$ , then

$$Pr_a(Re) = Pr_b(Re); \quad \text{when } Re \leq R_0 \text{ or } Re \geq R_3. \quad (13)$$

For  $\chi_a$ , let  $Pr_a = P_1$ ,  $Re_a = R_1$ , after ranking  $\mathbf{x}_j$ . For  $\chi_b$ , let  $Pr_b = P_2$ ,  $Re_b = R_2$ , after ranking  $\mathbf{x}_{j+1}$ . According to the analysis in Section 2.2, we have  $P_1 > P_2$  and  $R_1 > R_2$ . It can be find that  $Re_a$  jumps from  $R_0$  to  $R_1$  due to the contribution of  $\mathbf{x}_j$ , for convenience, we assume that  $Pr_a$  keeps unchanged when  $R_0 < Re_a < R_1$ . Under the assumption, we have two precision/recall graphs, *acd* and *abd* as shown in Fig.1, for  $\chi_a$  and  $\chi_b$ , respectively, when recall is in  $(R_0, R_3)$ . In the figure, the arrow means the precision jumps from the old value to the new value at the corresponding recall. Obviously, we have

$$Pr_a(Re) \geq Pr_b(Re); \quad \text{when } R_0 < Re < R_3. \quad (14)$$

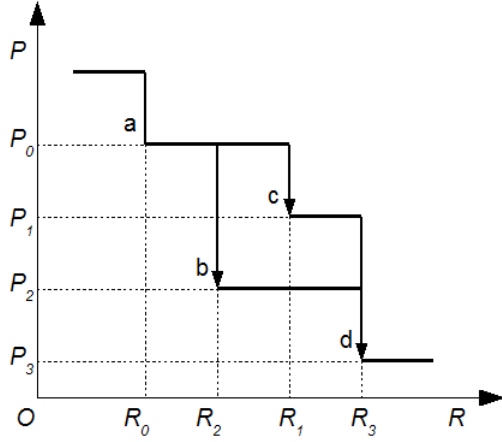


Figure 1: Case study

So from (13) and (14) we have

$$Pr_a(Re) \geq Pr_b(Re); \quad \forall Re. \quad (15)$$

That is, the performance of  $\chi_a$ ,  $Pr_a(Re)$ , is strongly better than the performance of  $\chi_b$ ,  $Pr_b(Re)$ .

Now we show the proposed ranking method can achieve optimal precision and recall.

If  $\chi_s$  is the sequence obtained by ranking  $\{\mathbf{x}_n\}$  according to their relative densities and  $\chi_o$  is any other sequence obtained by ranking  $\{\mathbf{x}_n\}$  with other method. Then the performance of  $\chi_s$ ,  $Pr_s(Re)$ , is strongly better than that of  $\chi_o$ ,  $Pr_o(Re)$ , that is,

$$Pr_s(Re) \geq Pr_o(Re); \quad \forall Re. \quad (16)$$

Without less of generality, assume  $f_{r,i} > f_{r,i+1}$ . We will use induction on  $n$ .  $n$  is the number of images.

*Base case:* When  $n = 2$ , we want to prove that the retrieval performance of the sequence obtained by ranking  $n$  images in accordance to their relative density is strongly better than that of any other sequences. Since it is a special case of ‘Case Study’ in which every sequence has 2 images, according to ‘Case Study’, it is true.

*Inductive step:* Suppose that for a given  $n \in N$ , the retrieval performance of the sequence obtained by ranking  $n$  images in accordance to their relative density is strongly better than that of any other sequences. (inductive hypothesis)

Our goal is to show that, the retrieval performance of the sequence obtained by ranking  $n + 1$  images in accordance to relative density is strong better than that of any other sequences. We will use the proof by contradiction to prove it as follows.

Let us assume that the negation of what we are trying to prove: that is there exists a sequence  $\chi_o$  that satisfies

$$Pr_o(Re) > Pr_s(Re); \quad \exists Re, \quad (17)$$

where  $Pr_o(Re)$  is the retrieval performance of the sequence

$$\chi_o := (\mathbf{x}_{r_1}, \dots, \mathbf{x}_{r_{n+1}}); \quad r_i \in N, \quad (18)$$

and  $Pr_s(Re)$  is the performance of the sequence  $\chi_s$  obtained by ranking images in accordance to their relative densities.

$$\chi_s := (\mathbf{x}_1, \dots, \mathbf{x}_{n+1}). \quad (19)$$

Considering  $\mathbf{x}_{r_{n+1}}$ , there are three cases:  $\mathbf{x}_{r_{n+1}} = \mathbf{x}_1$ ,  $\mathbf{x}_{r_{n+1}} = \mathbf{x}_i$  ( $1 < i \leq n$ ) and  $\mathbf{x}_{r_{n+1}} = \mathbf{x}_{n+1}$ .

Case 1: suppose  $\mathbf{x}_{r_{n+1}} = \mathbf{x}_1$ . We construct a sequence  $T$ .

$$T := (\mathbf{x}_2, \mathbf{x}_3, \dots, \mathbf{x}_{n+1}, \mathbf{x}_1). \quad (20)$$

To compare the retrieval performance of  $\chi_o$  and  $T$ , it needs only to compare the retrieval performance in the first  $n$  images since the last images in both sequences are  $\mathbf{x}_1$ . According to the inductive hypothesis, the retrieval performance of  $T$  is strongly better than that of  $\chi_o$ . We then construct another sequence  $T'$ .

$$T' := (\mathbf{x}_2, \mathbf{x}_1, \mathbf{x}_3, \dots, \mathbf{x}_{n+1}). \quad (21)$$

To compare the retrieval performances of  $T'$  and  $T$ , it needs only to compare the performance in the last  $n$  images since the first images in both sequences are  $\mathbf{x}_2$ . According to the inductive hypothesis, the retrieval performance of  $T'$  is strongly better than that of  $T$ . Due to the transitivity property, the retrieval performance of  $T'$  is strongly better than that of  $\chi_o$ .

To compare the performances of  $T'$  and  $\chi_s$ , it needs only to compare the performance in the first  $n$  images since the last images in both sequences are  $\mathbf{x}_{n+1}$ . According to the inductive hypothesis, the retrieval performance of  $\chi_s$  is strongly better than that of  $T'$ . Due to the transitivity property, the performance of  $\chi_s$  is strongly better than that of  $\chi_o$ .

Case 2: suppose  $\mathbf{x}_{r_{n+1}} = \mathbf{x}_i$ ,  $1 < i \leq n$ . Similarly, we construct two sequences  $T$  and  $T'$ ,

$$T := (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{i-1}, \mathbf{x}_{i+1}, \dots, \mathbf{x}_{n+1}, \mathbf{x}_i) \quad (22)$$

and

$$T' := (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{i-1}, \mathbf{x}_i, \mathbf{x}_{i+1}, \dots, \mathbf{x}_i, \mathbf{x}_{n+1}). \quad (23)$$

Then, according to the inductive hypothesis, we have the following relationships. The retrieval performance of  $T$  is strongly better than that of  $\chi_o$ . The retrieval performance of  $T'$  is strongly better than that of  $T$ . The retrieval performance of  $\chi_s$  is strongly better than that of  $T'$ . Due to the transitivity property, the retrieval performance of  $\chi_s$  is strongly better than that of  $\chi_o$ .

Case 3: suppose  $\mathbf{x}_{r_{n+1}} = \mathbf{x}_{n+1}$ . To compare the retrieval performances of  $\chi_o$  and  $\chi_s$ , it needs only to compare the retrieval performance in the first  $n$  images since the last images in both sequences are  $\mathbf{x}_{n+1}$ . According to the

inductive hypothesis, the performance of  $\chi_s$  is strongly better than that of  $\chi_o$ .

So the retrieval performance of  $\chi_s$  is strongly better than that of  $\chi_o$  in all three cases. That is

$$Pr_o(Re) \leq Pr_s(Re); \quad \forall Re. \quad (24)$$

But that contradicts our assumptions. So our assumptions false. So the retrieval performance of the sequence obtained by ranking  $n + 1$  images in accordance to their relative densities is strongly better than that of any other sequences.

Case study reveals that, if one image sequence is obtained using the proposed ranking method, changing the order of any two images will result in the retrieval performance degradation. It is the preparation for the following informal proof.

The informal proof shows that the retrieval performance of the proposed ranking method is better than that of any other ranking method when the two ranking method using the same feature distance metric. That is the proposed ranking method can achieve the optimal precision and recall.

## 4. Evaluation experiments

Experiments are conducted to evaluate the retrieval performance of the proposed ranking method with comparisons with conventional ranking methods. The image classes should be ranked using the method presented in [7]. Our task is to compute the relative density of images in class for ranking images. For convenience, we directly use the distances of ground truth images and the distances of all images in collection to estimate the distribution of relative density, then rank all images in collection in accordance of their relative density.

### 4.1. The Experiments

The IAPR TC-12 benchmark image collection (Image-CLEF2006) [13] is used in the experiments. It contains 20,000 photographic images. Based on the queries and their ground truth sets defined in the CLEF Cross-language Image Track 2006, we build up 20 ground truth sets for our experiments. Each ground truth set consists of about 40 ground truth images. Five standardized MPEG-7 visual descriptors [3] are used in the experiments including the Dominant Color Descriptor (DCD), the Color Layout Descriptor (CLD), the Color Structure Descriptor (CSD), the Edge Histogram Descriptor (EHD) and the Homogeneous Texture Descriptor (HTD).

In the experiments, the feature distances between query image and database image are computed using the functions recommended by MPEG-7. The overall distance is obtained by the simple equal weighting aggregation of multiple feature distances. Then the proposed ranking method is applied for ranking all images in collection, in which the histogram



(a) conventional method



(b) Proposed method

Figure 2: Retrieval results for the query “church”

techniques with 500 bins are used to estimate the densities. All collection images are put into bins in accordance to their normalized distances to the query.

### 4.2. The Results

In this subsection, we report the results of experiments that compare the retrieval performance of the proposed ranking method with the ranking method using feature distances on which most conventional image retrieval techniques are based.

To observe the difference of performances manifested in the ranked retrieval results, we present some image retrieval results. Fig.2(a) and (b) are 10 top ranked images from the conventional method and the proposed schemes for a special query, named “church”. The first image at the top-left in these figures is the query image. Fig.3(a) depicts the graphs of the average precision and recall for the query. It is clear that the proposed ranking method can retrieve more relevant images in semantic level than the conventional distanced based ranking method.

Fig.3(b) depicts the graphs of the average precision and recall in a single ground truth set, as an example. Fig.3(c) depicts the graphs of the average precision and recall over all ground truth sets. In both figures, the solid line represents the retrieval performance of the proposed ranking method and the dash line represents the ranking method using feature distances. In both cases, the results have shown that the proposed method has a significant improvement of retrieval performance over the ranking method using feature distances in the whole range of recalls.

## 5. Conclusion

Precision and recall are commonly used to evaluate the retrieval performance of CBIR systems. But conventional

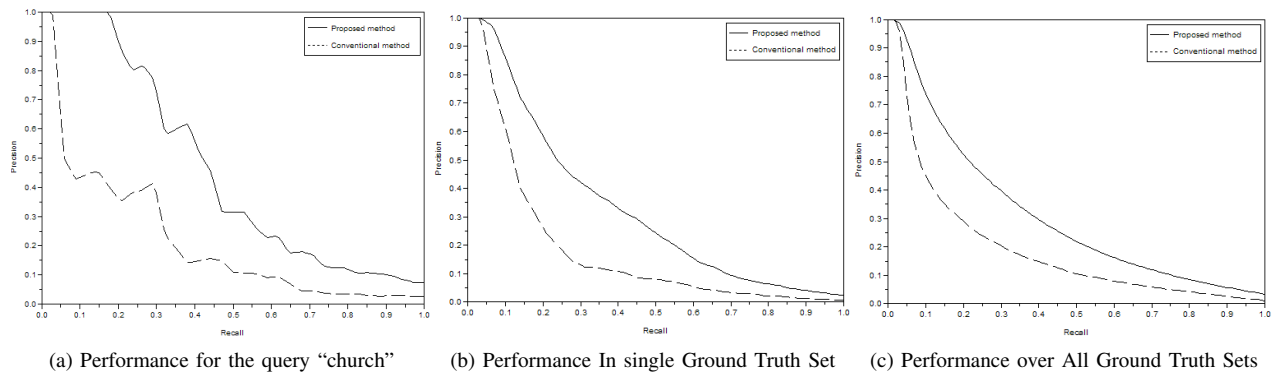


Figure 3: Evaluation of Retrieval Performance

retrieval methods or ranking methods, based on distance or probability, try to optimize average distances or retrieval errors other than precision/recall. The probability based ranking method proposed by N. Vasconcelos is for ranking image classes, but ranking images in the class is not considered. To address these issues, this paper proposed a new ranking method to improve the probability based ranking method by ranking images in the class under the precision/recall criteria. The method based on the relative density is derived from the maximum posterior probability (MAP) criterion, which is proved to be able to achieve optimal precision and recall. Experimental evaluation has shown that the proposed ranking method outperforms the ranking method using feature distances in all range of recall. Our future work will be focus on developing practical approaches to estimate the relative density.

## References

- [1] A. W. M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, "Content-based image retrieval at the end of the early years," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 12, pp. 1349–1380, Dec. 2000.
- [2] Y. Liu, D. S. Zhang, G. J. Lu, and W. Y. Ma, "A survey of content-based image retrieval with high-level semantics," *Pattern Recognition*, vol. 40, no. 1, pp. 262–282, Jan. 2007.
- [3] B. S. Manjunath, J. R. Ohm, V. V. Vasudevan, and A. Yamada, "Color and texture descriptors," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 11, no. 6, pp. 703–715, Jun. 2001.
- [4] Y. Rui, T. S. Huang, M. Ortega, and S. Mehrotra, "Relevance feedback: A power tool for interactive content-based image retrieval," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 8, no. 5, pp. 644–655, Sep. 1998.
- [5] S. Aksoy, R. Haralick, F. Cheikh, and M. Gabbouj, "A weighted distance approach to relevance feedback," in *Proc. Int. Conf. Pattern Recognition*, vol. 4, 2000, pp. 870–876.
- [6] A. Kushki, P. Androustos, K. N. Plataniotis, and A. N. Venetsanopoulos, "Retrieval of images from artistic repositories using a decision fusion framework," *IEEE Trans. Image Process.*, vol. 13, no. 3, pp. 277–292, Mar. 2004.
- [7] N. Vasconcelos, "On the efficient evaluation of probabilistic similarity functions for image retrieval," *IEEE Trans. Inf. Theory*, vol. 50, no. 7, pp. 1482–1495, Jul. 2004.
- [8] T. Westerveld and A. P. de Vries, "Generative probabilistic models for multimedia retrieval: query generation against document generation," *IEE Proc.-Vis. Image Signal Process.*, vol. 152, no. 6, pp. 852 – 858, Dec. 2005.
- [9] H. Muller, W. Muller, D. M. Squire, S. Marchand-Maillet, and T. Pun, "Performance evaluation in content-based image retrieval: overview and proposals," *Pattern Recognition Letters*, vol. 22, no. 5, pp. 593–601, Apr. 2001.
- [10] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*. New York: Wiley, 2001.
- [11] S. Aksoy and R. M. Haralick, "A classification framework for content-based image retrieval," in *Proc. IEEE Int. Conf. on Pattern Recognition*, vol. 2, Quebec, Canada, Aug. 2002, pp. 240–249.
- [12] R. P. W. Duin, "Classifiers in almost empty spaces," in *Proc. IEEE Int. Conf. on Pattern Recognition*, vol. 2, 2000, p. 17.
- [13] M. Grubinger, P. Clough, H. Muller, and T. Deselaers, "The iapr tc-12 benchmark: A new evaluation resource for visual information systems," in *Proc. Int. Workshop OntoImage'2006 Language Resources for Content-Based Image Retrieval, held in conjunction with LREC'06*, Genoa, Italy, May 2006, pp. 13–23.